Paper ET08

RADAR: Rapid AGILE Development and Automated Reporting for Clinical Data Standards (*Using R Shiny based Application*)

Neeraj Sharma, Syneos Heath, Gurugram, India
Dhirendra Kumar, Syneos Health, Gurugram, India

ABSTRACT

The pharmaceutical industry is under growing pressure to speed up clinical trial timelines without compromising regulatory compliance. Traditional waterfall development creates long cycles and struggles to adapt when study requirements change. We've developed a hybrid-Agile framework specifically for building R Shiny applications that generate CDISC-compliant SDTM and ADaM datasets, along with TFLs. Our six-phase approach—Plan, Design, Develop, Test, Review, Release—weaves together regulatory requirements, clinical validation processes, and automated quality controls. The framework takes advantage of what R Shiny does best: dynamic specification editing, real-time collaboration, and automated document generation for everything from SDRGs and ADRGs to define.XML files. In practice, we're seeing shorter development times, stronger quality control, faster regulatory submissions, and less manual work, all while maintaining data integrity and proper version control. This gives pharmaceutical companies a practical way to modernize their clinical data management without sacrificing the regulatory rigor they need.

INTRODUCTION

The landscape of clinical trial data management and statistical programming has changed considerably over the past ten years, driven by stricter regulatory requirements, bigger datasets, and the need for more sophisticated analytical tools. The traditional waterfall approach gives us structure and documentation, but it means long development cycles that struggle to keep up with how clinical trial requirements shift in practice. Waterfall handles regulatory compliance well—plan everything upfront, document extensively, validate at the end—but it's inflexible when requirements change. Agile embraces iterative development and welcomes changing requirements, but we must do that without losing regulatory rigor. CDISC standards—SDTM and ADAM specifically—have become the gold standard for FDA and EMA submissions, demanding precise data transformations, thorough validation, and extensive documentation. Implementing these standards in software means tracking data lineage through every transformation, maintaining traceability at every step, and validating continuously. What we've learned from successful Agile implementations in other regulated industries is that you need three things: get stakeholders involved early, document smartly, and take a risk-based approach to validation.

R Shiny has emerged as a powerful option for building interactive applications in clinical research, bringing together R's statistical capabilities with user-friendly web interfaces. It's become popular because statisticians can build sophisticated applications without being web developers. That said, there's a huge gap between building an exploratory Shiny app for internal review and building a production-ready application that will support regulatory submissions. The latter demands rigorous testing frameworks, comprehensive validation protocols, and deployment strategies that balance development agility with regulatory rigor. This paper lays out a comprehensive framework for implementing hybrid-Agile methodologies (RADAR) when developing R Shiny applications for clinical statistical programming.

FRAMEWORK DEVELOPMENT APPROACH

We developed our Agile framework (RADAR) through a systematic analysis of existing clinical programming workflows, regulatory requirements, and Agile best practices. The development process included:

- Requirements Analysis: Comprehensive review of CDISC standards, FDA guidance documents, and industry best practices for clinical data processing
- Stakeholder Consultation: Interacting with clinical data managers, biostatisticians, and statistical programmers to understand workflow requirements
- Technical Assessment: Evaluation of R Shiny capabilities and limitations in clinical GxP environment
- Framework Design: Integration/Modification of Agile principles as per statistical programming requirements

THE ADAPTED AGILE FRAMEWORK (RADAR: RAPID AGILE DEVELOPMENT AND AUTOMATED REPORTING)

The hybrid AGILE framework or RADAR includes the following phases for a statistical programmer:

PLAN:

- R Shiny application setup, project setup and initialization
- Plugging all necessary documents such as aCRF, Protocol, dummy metadata into the app

DESIGN:

- Designing the specifications for SDTM, ADaM and TFL generation
- Uploading TFL Shells and SAP as per study requirements.
- Deciding on number of datasets/TFLs has to be created in co-ordination with biostatistician

DEVELOP:

- Edit and update specifications as per requirements
- Generate datasets and TFLs as per plan and design phase
- Minimal coding required due to the capabilities of R-Shiny

TEST:

- Create QC datasets and TFLs to validate developer outputs
- Generating Pinnacle21 (P21) reports for compliance quality checks

REVIEW:

- Special portal access and provisions are introduced for the QA (Quality Assurance) team to access data and validate quality using pre-defined tools.
- Generate all documents (SDRG, ADRG, define.XML) automatically with minimal editing from the R-Shiny app.

RELEASE:

- Achieve final labeling and production status for final production or IA (Interim Analysis) with just a few clicks.

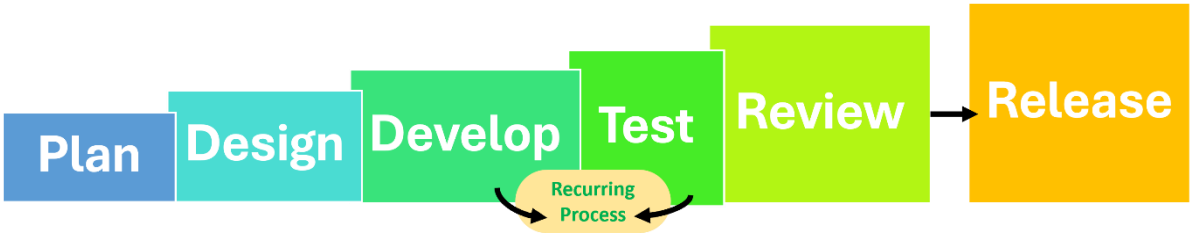

Figure1: showing different phases of AGILE methodology adopted for RADAR

SPRINT STRUCTURE AND TIMELINE

Our framework employs a modified sprint structure optimized for statistical programming requirements:
Sprint Duration: 2-3 weeks, aligned with clinical milestone schedules rather than arbitrary time boundaries.

Sprint 0 (Foundation Sprint – Happens before FPFV):

- Environment setup and configuration management
- Regulatory requirements gathering and creating aCRF
- CDISC standards review and mapping (Creating dummy metadata)
- Risk assessment and mitigation planning

Development Sprints (After the raw data has been received by statistical programmers):

- Feature development with continuous clinical validation
- Creating E2E CDISC compliant datasets (SDTM and ADaM) and TFLs
- Iterative testing with QA
- Reviewing raw data for data issues and consistency checks

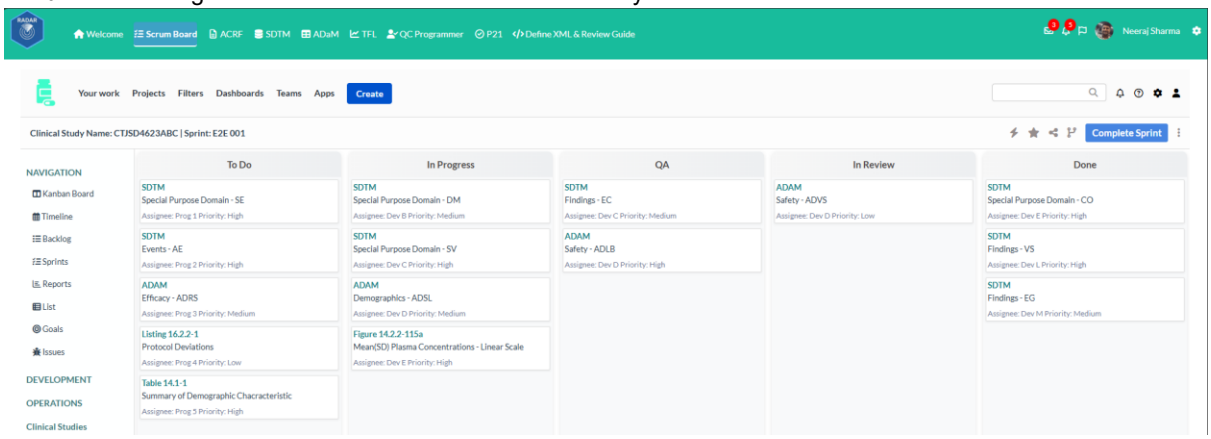

Figure2: showing example of a development sprint when using RADAR approach

Hardening Sprints (During the IAs – Interim Analysis and submissions):

- Comprehensive validation testing (QA final testing)
- Documentation completion (Review Guide and Define.XML)
- Regulatory compliance verification

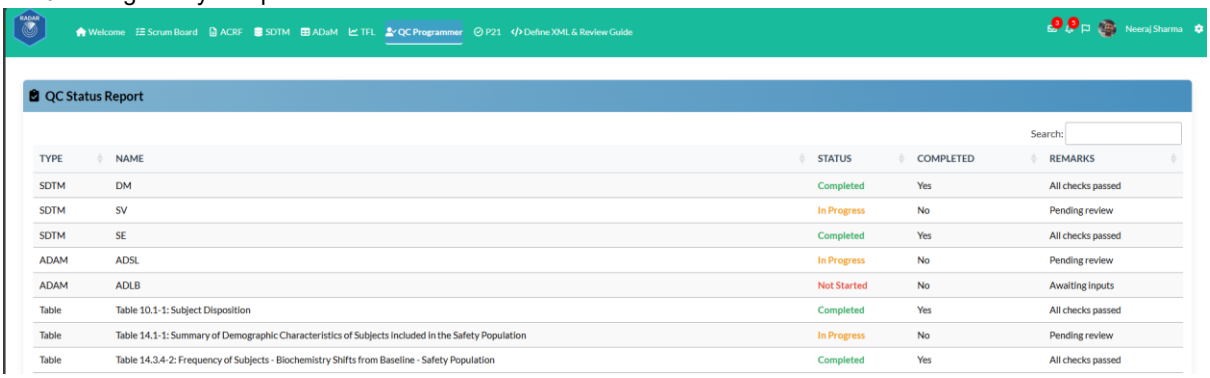

Figure3: showing example of QA board when using RADAR approach

ROLE DEFINITIONS AND RESPONSIBILITIES

Product Owner (Study Lead – Lead Statistical Programmer):

- Prioritizes features according to study timelines and regulatory milestones
- Validates outputs against clinical and statistical requirements

Scrum Master (Technical Lead):

- Statistical Programmer with reasonable regulatory experience
- Facilitates Agile ceremonies adapted for clinical context
- Ensures compliance with validation and documentation requirements

Development Team:

- R/Shiny developers with clinical domain knowledge
- Statistical programmers specializing in CDISC standards
- Quality assurance analysts with regulatory experience

RISK MANAGEMENT AND REGULATORY CONSIDERATIONS

TECHNICAL RISK MITIGATION (FOR R SHINY APPLICATION):

- Data Quality: The functioning of the app should not affect/interfere with the data.
- Performance: Regular profiling and optimization of app
- Regulatory Changes: Flexible architecture designed to accommodate regular CDISC updates

CLINICAL RISK MANAGEMENT (FOR DATASETS AND TFLS):

- Data Integrity: Comprehensive audit trails and change control procedures
- Validation Gaps: Independent review processes and cross-validation (SAS vs R) approaches in initial phases
- Timeline Risks: Advance/Adaptive Agile planning aligned with regulatory submission schedules

IMPLEMENTATION CONSIDERATIONS: PERFORMANCE METRICS AND SUCCESS CRITERIA

DEVELOPMENT EFFICIENCY METRICS:

- Sprint velocity measured in completed studies for all users
- Time from feature request (study-specific requests) to regulatory-ready implementation
- Bug detection and resolution rates during development phases

CLINICAL EFFECTIVENESS METRICS:

- Data transformation accuracy and error rates
- Evaluate processing time for datasets and TFL creation
- User satisfaction and training effectiveness measures

TRAINING AND EDUCATION

Agile principles need to be adapted for the statistical programming context. Statistical programmers require technical training on R Shiny development to build these applications effectively.

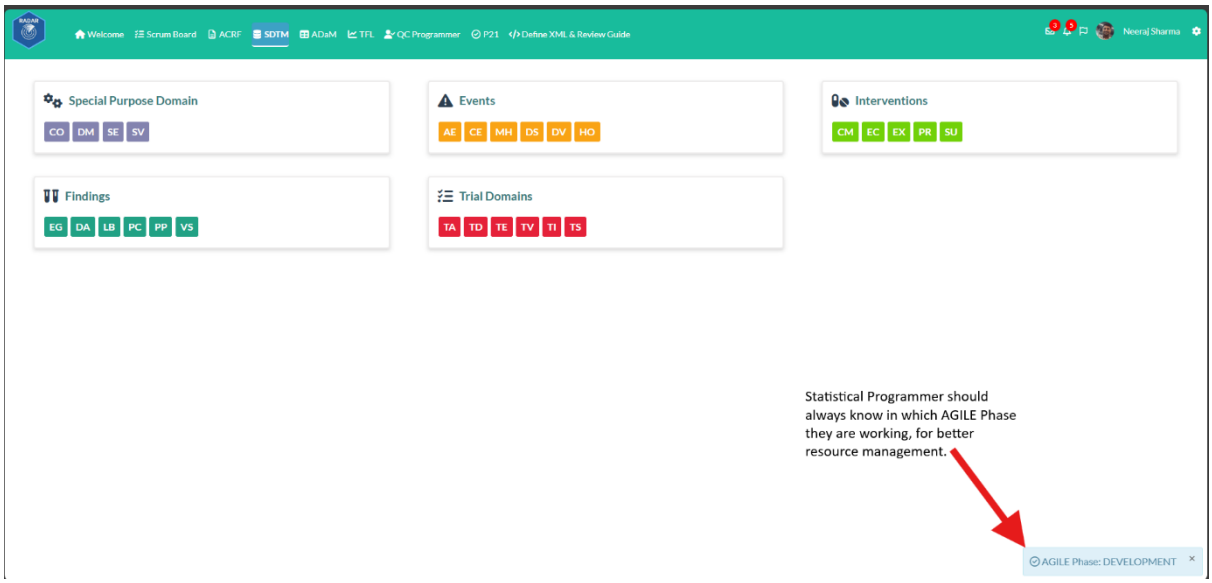

Figure4: showing AGILE phase notification for statistical programmers

STAKEHOLDER ENGAGEMENT

Regular communication with upstream teams like data management and downstream teams like medical writers is essential. Demonstrating value through pilot implementations helps build buy-in, while feedback incorporation and continuous improvement processes keep the framework evolving.

PROPOSED CASE STUDY FOR IMPLEMENTATION OF RADAR METHODOLOGY

To validate the framework's effectiveness, we implemented it in a mid-sized pharmaceutical company for the development of a comprehensive CDISC processing application. The project involved creating an R Shiny application for processing Phase III clinical trial data across multiple therapeutic areas.

PROJECT SCOPE

- 15 SDTM domains with complex derivation logic
- 8 ADAM datasets including population and efficacy analyses
- 50+ regulatory tables, figures, and listings
- Integration with existing clinical data management systems

IMPLEMENTATION TIMELINE

- 6 months total development time (compared to 12 months estimated for traditional approach)
- 12 two-week sprints with continuous validation
- 2 hardening sprints for final validation and documentation

TEAM SIZE AND COMPOSITION

Optimal results can be achieved with cross-functional team of 5-8 statistical programmers including both R and domain knowledge expertise.

ADOPTION STRATEGY

We recommend the following approach for organizations considering adoption:

- Pilot Implementation: Begin with a small, well-defined already validated clinical trial study to validate the approach to define the quality of outputs and team structure.
- Training Investment: Provide comprehensive training on both Agile principles and R-programming best practices as per CDISC/regulatory standards.
- Tool Standardization: Establish standardized development environments (Frozen R environments) and tool chains (backup R Shiny development team) to support consistent implementation.
- Metrics Collection: Implement comprehensive metrics collection to demonstrate value and guide continuous improvement using interactive R Shiny dashboard which can ingest real-time data from the pilot study as the study progress.

BENEFITS OF RADAR ADAPTATION IN CLINICAL PROGRAMMING

The implementation of our adapted Agile framework demonstrated several significant advantages over traditional waterfall approaches:

- Improved Stakeholder Engagement: Regular sprint reviews and continuous clinical validation ensured that developed features met actual user needs rather than theoretical requirements. This resulted in higher user adoption rates and reduced training requirements.
- Enhanced Quality Through Iterative Validation: The continuous validation approach allowed for early detection and correction of data quality issues, resulting in more robust final outputs and reduced regulatory review cycles.
- Flexibility in Requirement Changes: The Agile approach accommodated protocol amendments and regulatory guidance changes more effectively than traditional approaches, reducing project risks and timeline impacts.

FUTURE DIRECTIONS AND TECHNOLOGY EVOLUTION

The rapid evolution of R packages for clinical programming suggests several areas for framework enhancement:

- Machine Learning Integration: Incorporation of automated data quality checks and anomaly detection capabilities.
- Cloud-Native Development: Adaptation of the framework for cloud-based development (Posit cloud, Domino etc.) and deployment environments.
- Advanced Visualization: Integration of modern data visualization techniques (plotly, ggplot2) for enhanced regulatory reporting.

CONCLUSION

This paper presents a comprehensive framework for implementing hybrid Agile methodology (RADAR) in R Shiny application development for clinical statistical programming. The framework successfully addresses the unique challenges of developing regulatory-compliant applications while maintaining the benefits of iterative development and continuous stakeholder engagement. We can summaries the whole approach in the following points:

- Practical Framework: A detailed, implementable framework that balances Agile principles with regulatory requirements and providing CDISC compliant outputs
- Validation Strategy: Comprehensive approaches to testing and validation that address both technical and clinical requirements
- Implementation Guidance: Practical recommendations for organizational adoption and change management
- Performance Metrics: Demonstrated improvements in development efficiency and output quality

ACKNOWLEDGMENT

We want to thank Tasnima Abedin (for helping me to write the abstract) and Chanchal Bhalla (for helping me to develop R Shiny application as per RADAR methodology) for their constant support and pushing me to publish our idea.

CONTACT INFORMATION

Neeraj Sharma
Syneos Health
neeraj.sharma@syneoshealth.com

Dhirendra Kumar
Syneos Health
dhirendra.kumar@syneoshealth.com